

# Dependence with complete connections and the Gauss-Kuzmin theorem for $N$ -continued fractions

Dan Lascu\*

October 14, 2015

## Abstract

We consider a family  $\{T_N : N \geq 1\}$  of interval maps as generalizations of the Gauss transformation. For the continued fraction expansion arising from  $T_N$ , we solve its Gauss-Kuzmin-type problem by applying the theory of random systems with complete connections by Iosifescu.

**Mathematics Subject Classifications (2010).** 11J70, 11K50

**Key words:** continued fractions, invariant measure, random system with complete connections

## 1 Introduction

The purpose of this paper is to prove a GaussKuzmin type problem for  $N$ -continued fraction expansions introduced by Burger et al. [3]. In order to solve the problem, we apply the theory of random systems with complete connections by Iosifescu [9].

Fix an integer  $N \geq 1$ . In this paper, we consider a generalization of the Gauss transformation, i.e.,

$$T_N(x) := \begin{cases} \frac{N}{x} - \left\lfloor \frac{N}{x} \right\rfloor, & x \in I := [0, 1], x \neq 0, \\ 0, & x = 0 \end{cases} \quad (1.1)$$

where  $\lfloor \cdot \rfloor$  denotes the floor (or entire) function.

---

\*e-mail: lascudan@gmail.com.

The generalized Gauss measure  $G_N(A) = \frac{1}{\log\{(N+1)/N\}} \int_A \frac{dx}{x+N}$ ,  $A \in \mathcal{B}_I = \sigma$ -algebra of all Borel subsets of  $[0, 1]$  is  $T_N$ -invariant, i.e.,  $G_N(T_N^{-1}(A)) = G_N(A)$  for any  $A \in \mathcal{B}_I$ . Define  $a_1(x) = \lfloor \frac{N}{x} \rfloor$ ,  $x \in (0, 1]$ ,  $a_1(0) = \infty$ , and  $a_n(x) = a_1(T_N^{n-1}(x))$ ,  $x \in I$ ,  $n \in \mathbb{N}_+ := \{1, 2, \dots\}$ , with  $T_N^0(x) = x$ . By the very definitions, Burger et al. proved in [3] that any irrational  $0 < x < 1$  can be written in the form

$$x = \frac{N}{a_1 + \frac{N}{a_2 + \frac{N}{a_3 + \ddots}}} := [a_1, a_2, a_3, \dots]_N \quad (1.2)$$

where  $a_n$ 's are non-negative integers. We will call (1.2) the *N-continued fraction expansion of x*. In [4], Dajani et al. proved that  $(I, \mathcal{B}_I, G_N, T_N)$  is an ergodic dynamical system.

The Perron-Frobenius operator of  $T_N$  under a probability measure  $\mu$  on  $\mathcal{B}_I$  such that  $\mu(T_N^{-1}(A)) = 0$  whenever  $\mu(A) = 0$  is defined as the bounded linear operator  $U$  on the Banach space  $L^1(I, \mu) := \{f : I \rightarrow \mathbb{C} : \int_I |f| d\mu < \infty\}$  such that the following holds:

$$\int_A Uf d\mu = \int_{T_N^{-1}(A)} f d\mu \quad \text{for all } A \in \mathcal{B}_I, f \in L^1(I, \mu). \quad (1.3)$$

In particular, the Perron-Frobenius operator of  $T_N$  under  $G_N$  is

$$\{Uf\}(x) = \sum_{i \geq N} V_{N,i}(x) f\left(\frac{N}{x+i}\right), \quad f \in L^1(I, G_N) \quad (1.4)$$

where  $V_{N,i} := \frac{x+N}{(x+i)(x+i+1)}$  for any  $i \geq N$  and  $n \in \mathbb{N}_+$  [17]. An important fact is that for any probability measure  $\mu$  on  $\mathcal{B}_I$  such that  $\mu \ll \lambda$ , where  $\lambda$  is the Lebesgue measure on  $\mathcal{B}_I$ , we have

$$\mu(T_N^{-n}(A)) = \int_A \{U^n f\}(x) dG_N(x) \quad (1.5)$$

where  $f(x) := (\log(\frac{N+1}{N}))(x+N)h(x)$  for  $x \in I$  [17].

The problem of finding the asymptotic of  $T_N^{-n}(A)$  as  $n \rightarrow \infty$  represents the Gauss-Kuzmin-type problem for  $N$ -continued fraction expansions.

**Theorem 1.1.** (A Gauss-Kuzmin theorem for  $T_N$ ) *Fix an integer  $N \geq 1$  and let  $(I, \mathcal{B}_I, T_N)$  be as above.*

(i) For a probability measure  $\mu$  on  $(I, \mathcal{B}_I)$ , let the assumption (A) as follows:

(A)  $\mu$  is non-atomic and has a Riemann-integrable density.

Then for any probability measure  $\mu$  which satisfies (A), the following holds:

$$\lim_{n \rightarrow \infty} \mu(G_N^n < x) = \frac{1}{\log\{(N+1)/N\}} \log \frac{x+N}{N}, \quad x \in I. \quad (1.6)$$

(ii) In addition to assumptions of  $\mu$  in (i), if the density of  $I \ni x \mapsto \mu([0, x])$  is Lipschitz continuous, then there exist two positive constants  $q < 1$  and  $\ell$  such that for any  $x \in I$  and  $n \geq 1$ , there exists  $\theta$  with  $|\theta| \leq \ell$ , the following holds:

$$\mu(G_N^n < x) = \frac{1 + \theta q^n}{\log\{(N+1)/N\}} \log \frac{x+N}{N} \quad (1.7)$$

where  $\theta := \theta(N, \mu, n, x)$ . As a consequence, the  $n$ -th error term  $e_n(N, \mu; x)$  of the Gauss-Kuzmin problem is obtained as follows:

$$e_n(N, \mu; x) = \frac{\theta q^n}{\log\{(N+1)/N\}} \log \frac{x+N}{N}. \quad (1.8)$$

The paper is organized as follows. In Section 2, we recall random system with complete connections (=RSCC), and show examples and properties. In Section 3, we prove Theorem 1.1. By using the ergodic behavior of the RSCC associated with  $N$ -continued fraction expansion, we determine the limit of the sequence  $(\mu(G_N^n < x))_{n \geq 1}$  of distributions as  $n \rightarrow \infty$ .

## 2 Random systems with complete connections

In this section, we introduce random system with complete connections (=RSCC) and show its properties.

The first explicit formal definition of the concept of dependence with complete connections was given by Onicescu and Mihoc [22]. It is a non-trivial extension of Markovian dependence theory, and it was also investigated by Doeblin and Fortet [5] and by Harris [7]. The concept of random system with complete connections was defined by Iosifescu [8].

Examples of RSCC are urns models [22, 10], stochastic learning processes [21, 10, 12], partially observed random chains [10], image coding [1], continued fraction expansions [9, 16, 23, 24, 25] and others.

An RSCC is often called an *iterated function system with place-dependent probabilities* or simply an *iterated function system (=IFS)* [2]. Applications of IFS to continued fractions can be found in [15, 19]. For more detail, see [9, 10, 11].

## 2.1 Definitions and their extensions

In this subsection, we introduce the definition of random system with complete connections which is restricted to a smaller class than the original. Next we extend domains of functions in such a system.

**Definition 2.1.** [8, 9, 21] *A random system with complete connections (=RSCC) is a quadruple*

$$\{(W, \mathcal{W}), X, u, P\} \quad (2.1)$$

where

- (i)  $(W, \mathcal{W})$  is a measurable space and  $X$  is a non-empty set;
- (ii)  $u : W \times X \rightarrow W$  is a measurable function with respect to  $\mathcal{W}$  for each element in  $X$ ;
- (iii)  $P : W \times X \rightarrow [0, 1]$  satisfies that  $\sum_{x \in X} P(w, x) = 1$  for each  $w \in W$ , and  $P(\cdot, x)$  is a measurable function on  $(W, \mathcal{W})$  for each  $x \in X$ .

For an RSCC in Definition 2.1, we call  $W$ ,  $X$ ,  $u$  and  $P$ , the *state space*, the *event space*, the *transition function* and the  $((W, X)$ -) *transition probability function*, respectively ([21], Section 1.2). The role of the function  $u$  is to change a state  $w \in W$  into the new state  $w' = u(w, x) \in W$  by an event  $x \in X$ :

$$W \ni w \xrightarrow{x} w' = u(w, x) \in W. \quad (2.2)$$

In this case,  $P(w, x)$  is regarded as the probability of the transition  $w \mapsto w'$  which depends on the information of both  $w$  and  $x$ .

**Remark 2.2.** (i) In the original definition of RSCC in [9],  $X$  is assumed as a measurable space  $(X, \mathcal{X})$ . In Definition 2.1, we treat only the case  $\mathcal{X} = \mathcal{P}(X)$  (= the power set of  $X$ ). A lot of examples in [9] satisfy this condition.

- (ii) An RSCC can be regarded as an automaton with output [20, 13]. A *Mealy machine*  $(W, X, \Delta, u, \lambda)$  consists of three sets  $W, X, \Delta$  and two maps  $u, \lambda$  where  $W$  is the *set of states*,  $X$  is the *input alphabet*,  $\Delta$  is

the *output alphabet*, and the *transition function*  $u : W \times X^* \rightarrow W$  and the *map of outputs*  $\lambda : W \times X^* \rightarrow \Delta^*$  which satisfy

$$\begin{cases} u(w, xa) = u(u(w, x), a), \\ \lambda(w, xa) = \lambda(w, x)\lambda(u(w, x), a), \end{cases} \quad x \in X^*, a \in X, w \in W \quad (2.3)$$

where  $X^*$  and  $\Delta^*$  denote free semigroups generated by  $X$  and  $\Delta$ , respectively. When it is an RSCC,  $\Delta = [0, 1]$  and  $\Delta^*$  is reduced to the  $[0, 1]$  with respect to the multiplication in  $(\mathbb{R}, \times)$ ,  $\pi : \Delta^* \rightarrow [0, 1]$ . The map  $\pi \circ \lambda$  is a transition probability function  $P$ . An example of Mealy machine as an RSCC will be shown in Example 2.5. The similarity can be explained as the reason that initial applications of RSCC were related to learning processes which may be understood as an algorithmic study of such systems.

For a given RSCC  $\{(W, W), X, u, P\}$ , we extend domains of both  $u$  and  $P$  as follows.

- (i) In order to extend the domain of  $u$ , we prepare the free semigroup  $X^*$  generated by  $X$  as follows: In Definition 2.1, we regard  $X$  as a set of letters, and treat  $X^n$  as the set of all strings of length  $n \geq 1$ . We write an element  $(x_1, \dots, x_n) \in X^n$  as  $x_1 \cdots x_n$ . Then the set  $X^*$  of all strings with the empty letter  $\varepsilon$  is a semigroup with the concatenation as its product, and the unit of  $X^*$  is  $\varepsilon$ . The semigroup  $X^*$  is called the *free semigroup generated by  $X$*  [18].

With respect to the semigroup  $X^*$ , the function  $u$  induces a unique right action of  $X^*$  on  $W$  which is denoted by the same symbol  $u$  here:

$$u : W \times X^* \rightarrow W; \quad (w, x) \mapsto wx := u(w, x). \quad (2.4)$$

For example,  $w(xx')$  is defined as  $(wx)x' = u(u(w, x), x')$  for  $w \in W$ ,  $x \in X^*$  and  $x' \in X$ . For  $x = x_1 \cdots x_n \in X^n$ , we can write  $wx := w(x_1 \cdots x_n)$  for  $w \in W$ .

- (ii) Let  $\mathcal{P}(X)$  denote the power set of  $X$ . The domain  $W \times X$  of  $P$  is extended to  $W \times \mathcal{P}(X)$  as follows:

$$P(w, A) := \begin{cases} \sum_{x \in A} P(w, x), & A \neq \emptyset, \\ 0, & A = \emptyset, \end{cases} \quad (w, A) \in W \times \mathcal{P}(X). \quad (2.5)$$

By Definition 2.1(iii), we see that  $(X, \mathcal{P}(X), P(w, \cdot))$  is a probability space for each  $w \in W$ .

As a generalization of  $P$ , we define  $P_r : W \times X^r \rightarrow [0, 1]$  for each  $r \geq 1$  by

$$P_1 := P, \quad P_r(w, xx') := P_{r-1}(w, x)P(wx, x'), \quad (2.6)$$

with  $w \in W, x \in X^{r-1}, x \in X, r \geq 2$ , where the notation in (2.4) is used. Then we see that  $P_r(w, x)$  means the transition probability from  $w$  to  $w \cdot x$  along the path  $w \rightarrow wx_1 \rightarrow wx_1x_2 \rightarrow \cdots \rightarrow wx_1 \cdots x_r$  with length  $r$ . For  $A \subset X^r$ , we also define

$$P_r(w, A) := \begin{cases} \sum_{x \in A} P_r(w, x), & A \neq \emptyset, \\ 0, & A = \emptyset. \end{cases} \quad (2.7)$$

By definition,  $P_r(w, A)$  is the summation of all paths from  $w$  to  $w \cdot x$  for  $x \in A$ . For example, we can verify that  $P_r(w, A \times X) = P_{r-1}(w, A)$  for any  $(w, A) \in W \times \mathcal{P}(X^{r-1})$  by Definition 2.1(iii) and (2.6). Especially,  $P_r(w, X^r) = 1$ . Hence  $(X^r, \mathcal{P}(X^r), P_r(w, \cdot))$  is also a probability space for each  $w \in W$  and  $r \geq 1$ .

## 2.2 Transition probability functions defined on the square of the state space

Let  $\{(W, \mathcal{W}), X, u, P\}$  be as in Definition 2.1. Recall that  $P$  is the transition probability function with the domain  $W \times \mathcal{P}(X)$ . We define new transition probability functions with the domain  $W \times \mathcal{W}$  in this subsection.

### 2.2.1 Space of paths

For  $w, w' \in W$ , define the subset  $X(w, w')$  of  $X$  by

$$X(w, w') := \{x \in X : u(w, x) = w'\}. \quad (2.8)$$

Then  $X(w, w')$  can be identified with the set of all paths from  $w$  to  $w'$  with length 1. Remark that  $X(w, w')$  may be the empty set. By definition,  $X(w, w') \cap X(w, w'') = \emptyset$  when  $w' \neq w''$  and  $\coprod_{w' \in W} X(w, w') = X$ . From these properties, the following holds.

**Fact 2.3.** *For  $(w, B) \in W \times \mathcal{W}$ , let  $X(w, B) := \{x \in X : u(w, x) \in B\}$ . Then the following holds for each  $w \in W$ :*

(i)  $X(w, B) \cap X(w, B') = \emptyset$  when  $B \cap B' = \emptyset$ .

(ii)  $X(w, B) = \bigcup_{w' \in B} X(w, w')$ .

(iii)  $X(w, B) \subset X(w, B')$  when  $B \subset B'$ .

(iv)  $X(w, B) \cup X(w, B') = X(w, B \cup B')$ .

In [9], the symbol  $X(w, B)$  is written as  $B_w$ .

### 2.2.2 Transition probability functions defined on the square of the state space

By using the extension of  $P$  in (2.5) and  $X(w, w')$  in (2.8), define the new function  $Q : W \times W \rightarrow [0, 1]$  by

$$Q(w, w') := P(w, X(w, w')), \quad (w, w') \in W \times W. \quad (2.9)$$

This means the total sum of transition probabilities from  $w$  to  $w'$  by all possible event  $x \in X$ . We extend the domain of  $Q$  to  $W \times \mathcal{W}$  as follows ([9], (1.1.11)):

$$Q(w, B) := P(w, X(w, B)) = \sum_{x \in X(w, B)} P(w, x), \quad (w, B) \in W \times \mathcal{W}. \quad (2.10)$$

We call  $Q$  the  $((W, W)-)$  transition probability function associated with  $\{(W, \mathcal{W}), X, u, P\}$ . The number  $Q(w, B)$  means the probability such that a state  $w$  changes to a state belonging to  $B$  by measuring all possible (one-step) events. For  $w \in W$ , define the function  $Q_w$  as

$$Q_w : \mathcal{W} \rightarrow [0, 1]; \quad Q_w(B) := Q(w, B), \quad B \in \mathcal{W}. \quad (2.11)$$

By Fact 2.3,

$$0 \leq Q_w(B) \leq Q_w(W) = 1, \quad Q_w(B) + Q_w(B') = Q_w(B \cup B') \text{ when } B \cap B' = \emptyset. \quad (2.12)$$

Hence  $(W, \mathcal{W}, Q_w)$  is a probability space for each  $w \in W$ . Remark that  $Q_w$  can be defined on the whole of  $\mathcal{P}(W)$ , but we restrict  $Q_w$  on the given  $\sigma$ -algebra  $\mathcal{W} \subset \mathcal{P}(W)$  here. Thanks to  $Q_w$ , we can define the integration  $\int_B f(w') dQ_w(w')$  for  $B \in \mathcal{W}$  and a measurable function  $f$  on  $(W, \mathcal{W})$ .

For  $w, w' \in W$ , define the family  $\{Q^{(k)}(w, w') : k \geq 1\}$  recursively as

$$Q^{(1)}(w, w') := Q(w, w'), \quad Q^{(k)}(w, w') := \int_W dQ_w(w'') Q^{(k-1)}(w'', w'), \quad k \geq 2. \quad (2.13)$$

By definition,  $Q^{(k)}(w, w')$  is the expectation value of  $Q^{(k-1)}(\cdot, w')$  on the probability space  $(W, \mathcal{W}, Q_w)$ . We extend the domain of  $Q^{(k)}$  to  $W \times \mathcal{W}$  as follows:

$$Q^{(1)}(w, B) := Q(w, B), \quad Q^{(k)}(w, B) := \int_W dQ_w(w') Q^{(k-1)}(w', B), \quad k \geq 2. \quad (2.14)$$

We will say  $Q^{(k)}$  the *k-step transition probability function of the Markov chain* associated with  $\{(W, \mathcal{W}), X, u, P\}$ . We see that  $(W, \mathcal{W}, Q_w^{(k)})$  is also a probability space for each  $w \in W$  and  $k \geq 1$  where  $Q_w^{(k)} := Q^{(k)}(w, \cdot)$ . Furthermore, define

$$Q_n(w, B) := \frac{1}{n} \sum_{k=1}^n Q^{(k)}(w, B), \quad (w, B) \in W \times \mathcal{W}, \quad n \geq 1. \quad (2.15)$$

For example,  $Q_1 = Q$ . Then  $(W, \mathcal{W}, Q_{n,w})$  is also a probability space for each  $w \in W$  where  $Q_{n,w}(B) := Q_n(w, B)$  for  $B \in \mathcal{W}$ .

### 2.3 Examples

In this section, we shall give two examples of RSCC.

**Example 2.4.** We show the RSCC associated with regular continued fraction expansions. More precisely, it is the RSCC associated with the dynamical system  $(I, \mathcal{B}_I, \tau)$  where  $\tau$  is the Gauss transformation, defined as  $\tau = T_1$ , where  $T_N$  is as in (1.1). Define the RSCC  $\{(W, \mathcal{W}), X, u, P\}$  as  $W = [0, 1]$ ,  $\mathcal{W} = \mathcal{B}_{[0,1]}$ ,  $X = \mathbb{N}_+$ ,

$$u : W \times X \rightarrow W; \quad u(w, x) = \frac{1}{w+x}, \quad (2.16)$$

$$P : W \times X \rightarrow [0, 1]; \quad P(w, x) = \frac{w+1}{(w+x)(w+x+1)}. \quad (2.17)$$

By definition, we see that  $u(\cdot, x)$  is also a right inverse of  $\tau$  for each  $x \in X$ , that is,  $\tau(u(w, x)) = w$  for any  $w \in W$ . This shows that the dynamical system  $(I, \mathcal{B}_I, \tau)$  is encoded onto the RSCC  $\{(W, \mathcal{W}), X, u, P\}$ . About more details, see Sec.1.2 of [9].

**Example 2.5.** According to §3.1 of [13], we show a simple (but non-trivial) example of (finite) RSCC by using a finite automaton with input/output. Define the data  $\{(W, \mathcal{W}), X, u, P\}$  as follows:



- (i) For the 2-point set  $\{1, 2\}$ , let  $\mathcal{P}(\{1, 2\})$  denote its power set. Then  $(\{1, 2\}, \mathcal{P}(\{1, 2\}))$  is a measurable space. Define  $(W, \mathcal{W}) := (\{1, 2\}, \mathcal{P}(\{1, 2\}))$  and  $X := \{1, 2\}$ .
- (ii) Any function  $u : W \times X \rightarrow W$  is measurable with respect to  $(W, \mathcal{W})$  by definition. For example, let

$$u(i, j) := j, \quad i, j = 1, 2. \quad (2.18)$$

- (iii) Any transition probability function  $P : W \times X \rightarrow [0, 1]$  is uniquely defined by two real numbers  $P(i, 1) \in [0, 1]$  for  $i = 1, 2$ . For example, for  $0 \leq \alpha, \beta \leq 1$ , define  $P(1, 1) = \alpha$  and  $P(2, 1) = \beta$ . Then they define a unique transition probability function  $P$  as

$$P(1, 1) = \alpha, \quad P(1, 2) = 1 - \alpha, \quad P(2, 1) = \beta, \quad P(2, 2) = 1 - \beta. \quad (2.19)$$

For example, the value  $Q(1, 1)$  of  $Q$  in (2.9) is computed as follows:

$$Q(1, 1) = P(1, \{x \in \{1, 2\} : u(1, x) = 1\}) = P(1, 1) = \alpha. \quad (2.20)$$

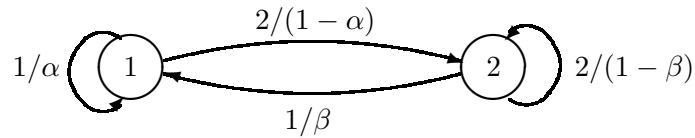
Remark that  $u$  in (2.18) is the case such that  $u(i, \cdot)$  does not depend on  $i \in W$  ([9], p15, (i)). As the case such that  $u(i, \cdot)$  depends on  $i \in W$ , we can define other  $u$ , for example,  $u(i, j) = i$  for  $i \in W, j \in X$ .

We can illustrate this example as a finite automaton with input/output (= Mealy machine [20]). Assume that  $\{(W, \mathcal{W}), X, u, P\}$  is as in (2.18) and (2.19). For this system, we draw the transition diagram (= an oriented graph with labeling edges) as follows:

- (1) The set of vertices is the state space  $W = \{1, 2\}$ .
- (2) For two vertices  $i, j \in W$  ( $i$  and  $j$  may be same), if there exists an event (=input signal)  $k \in X = \{1, 2\}$  such that  $u(i, k) = j$ , then draw the oriented edge from  $i$  to  $j$ . We write  $(i, k, j)$  as this edge here.
- (3) Write “ $k/P(i, k)$ ” as the label of an edge  $(i, k, j)$ .

According to these rules, the transition diagram of the RSCC is illustrated as Figure 2.6.

**Figure 2.6.**



For example, from (2.18),  $u(1, 1) = 1$ , and from (2.19),  $P(1, 1) = \alpha$ . Hence the label of the edge  $(1, 1, 1)$  is given as “ $1/\alpha$ ”. About a Markov chain related to this example, see Chap. 5 of [6]. About other examples of Mealy machine, see [13, 14].

## 2.4 Operators associated with an RSCC

In this subsection, we present the asymptotic and ergodic properties of operators associated with an RSCC. To state these results, we prepare definitions. We add the following assumption for all RSCC in this subsection:

**Assumption (B)** For an RSCC  $\{(W, \mathcal{W}), X, u, P\}$ ,  $W$  is a measurable subset of the measurable space  $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$  and  $\mathcal{W} = \mathcal{B}_W$ .

Thanks to Assumption (B), we can use the absolute value  $|\cdot|$  and the Lebesgue measure on  $W$  induced by  $\mathbb{R}$ , which will be necessary in this subsection.

Let  $L^\infty(W)$  denote the Banach space of all complex-valued bounded measurable functions defined on  $(W, \mathcal{W})$ . We define operators on  $L^\infty(W)$  associated with an RSCC  $\{(W, \mathcal{W}), X, u, P\}$  as follows.

**Definition 2.7.** (i) The transition operator  $U$  on  $L^\infty(W)$  is defined by

$$\{Uf\}(w) := \sum_{x \in X} P(w, x) f(u(w, x)), \quad f \in L^\infty(W), w \in W. \quad (2.21)$$

(ii) ([9], (3.1.7)) For  $n \geq 1$ , define the operator  $U_n$  on  $L^\infty(W)$  as

$$\{U_n f\}(w) := \int_W f(w') dQ_{n,w}(w'), \quad f \in L^\infty(W), w \in W \quad (2.22)$$

where  $Q_{n,w} := Q_n(w, \cdot)$  is as in (2.15).

**Remark 2.8.** (i) Let  $U$  be as in (2.21). For each  $w \in W$ ,  $\{U(f)\}(w)$  is the expectation value of  $(f \circ u)(w, \cdot)$  with respect to the probability space  $(X, \mathcal{P}(X), P(w, \cdot))$ . For example, if  $f$  is the characteristic function of  $B \in \mathcal{W}$ , then we see  $\{Uf\}(w) = Q(w, B)$ . For each  $n \geq 1$ ,  $\{U^n f\}(w) = \sum_{x \in X^n} P_n(w, x) f(wx)$  where  $U^n$  denotes the  $n$ -th iterate of  $U$  and  $P_n$  is as in (2.6) for  $r = n$ .

- (ii) Let  $U_n$  be as in (2.22). Remark that  $U_n$  is well-defined on  $L^\infty(W)$  because  $Q_{n,w}$  is a probability measure. For example, if  $f$  is the characteristic function of  $B \in \mathcal{W}$ , then  $\{U_1 f\}(w) = Q(w, B)$ . Since

$$Q_{n,w}(w') = Q_n(w, w') = \frac{1}{n} \sum_{k=1}^n Q^{(k)}(w, w'), \quad (2.23)$$

we see that

$$\{U_n f\}(w) = \frac{1}{n} \sum_{k=1}^n \int_W f(w') dQ_w^{(k)}(w'). \quad (2.24)$$

Next, let  $L(W)$  denote the Banach space of all complex-valued Lipschitz continuous functions on  $W$  with the following norm  $\|\cdot\|_L$ :

$$\|f\|_L := \sup_{w \in W} |f(w)| + \sup_{w' \neq w''} \frac{|f(w') - f(w'')|}{|w' - w''|}, \quad f \in L(W). \quad (2.25)$$

Remark that we use Assumption (B) here. By definition,  $L(W) \subset L^\infty(W)$ .

According to [9, 21], we introduce several characterizations of the operator  $U$  in (2.21) as follows.

**Definition 2.9.** ([9], Definition 3.1.4, [21], Definition 2.1) Let  $W, U, U_n, L(W)$  be as in (2.1), (2.21), (2.22) and (2.25), respectively. We consider restrictions of  $U, U_n$  on  $L(W)$  as follows.

- (i)  $U$  is ordered if there exists a bounded linear operator  $S$  on  $L(W)$  such that

$$\lim_{n \rightarrow \infty} \|U_n f - S f\|_L = 0, \quad f \in L(W). \quad (2.26)$$

- (ii)  $U$  is aperiodic if there exists a bounded linear operator  $S'$  on  $L(W)$  such that

$$\lim_{n \rightarrow \infty} \|U^n f - S' f\|_L = 0, \quad f \in L(W), \quad (2.27)$$

where  $U^n$  is the  $n$ -th iterate of  $U$  for  $n \geq 1$ .

- (iii)  $U$  is ergodic with respect to  $L(W)$  if  $U$  is ordered and the rank of  $S$  in (2.26) is 1.

- (iv)  $U$  is regular with respect to  $L(W)$  if  $U$  is ergodic and aperiodic.

- (v) The Markov chain corresponding to  $U$  is regular if  $U$  is regular with respect to  $L(W)$ .

**Remark 2.10.** (i) Definition 2.9(i) and (ii) mean that sequences  $(U_n)$  and  $(U^n)$  converge to operators  $S$  and  $S'$ , respectively, with respect to the strong operator topology on  $L(W)$ .

(ii) Under the Assumption (B),  $W$  is separable. In addition, if  $W$  is complete and  $U$  is orderly, then there exists probability measures  $\{Q_w^\infty : w \in W\}$  on  $(W, \mathcal{W})$  such that

$$\{Sf\}(w) = \int_W dQ_w^\infty(w') f(w'), \quad f \in L(W), w \in W \quad (2.28)$$

by Theorem 3.1.24 of [9] where  $S$  is as in Definition 2.9(i).

(iii) In addition to the assumption in (ii), if  $U$  is ergodic with respect to  $L(W)$ , then  $Sf$  in (2.28) is a constant function on  $W$  for any  $f \in L(W)$ . Therefore the operator  $S$  is identified with a bounded linear functional on  $L(W)$ ,  $S : L(W) \rightarrow \mathbb{C}$ . Then (2.28) is rewritten as

$$S : L(W) \rightarrow \mathbb{C}; \quad Sf = \int_W dQ^\infty(w') f(w'), \quad f \in L(W) \quad (2.29)$$

for some probability measure  $Q^\infty$  on  $(W, \mathcal{W})$ .

**Definition 2.11.** ([9], Definition 3.1.15) Under the Assumption (B),  $\{(W, \mathcal{W}), X, u, P\}$  is an RSCC with contraction if the following conditions are satisfied:

(i)  $r_1 < \infty$ ,

(ii)  $r_\ell < 1$  for some  $\ell \geq 1$ , and

(iii)  $R < \infty$

where  $(r_k)$  and  $R$  are defined as

$$r_k := \sup_{w' \neq w''} \sum_{x \in X^k} P(w', x) \frac{|w'x - w''x|}{|w' - w''|}, \quad k \geq 1, \quad (2.30)$$

$$R := \sup_{A \subset X} \sup_{w' \neq w''} \frac{|P(w', A) - P(w'', A)|}{|w' - w''|}. \quad (2.31)$$

Remark that we use the Assumption (B) for  $|\cdot|$  in Definition 2.11.

When  $k = 1$ , (2.30) is rewritten by using  $u$  as follows:

$$r_1 = \sup_{w' \neq w''} \sum_{x \in X} P(w', x) \frac{|u(w', x) - u(w'', x)|}{|w' - w''|}. \quad (2.32)$$

If  $\sup_{w' \neq w''} |u(w', x) - u(w'', x)|/|w' - w''| < 1$ , then  $r_1 < 1$ . In this case, assumptions (i) and (ii) in Definition 2.11 are satisfied.

**Theorem 2.12.** *Under the Assumption (B), let  $\{(W, \mathcal{W}), X, u, P\}$  be an RSCC with contraction. For  $Q^{(n)}$  in (2.13), define  $(\sigma_n)$  by*

$$\sigma_n(w) := \text{supp } Q^{(n)}(w, \cdot), \quad w \in W \quad (2.33)$$

where  $\text{supp } \mu$  denotes the support of a measure  $\mu$ . Assume that  $W$  is compact. Then the following holds.

- (i) *The Markov chain associated with the RSCC is regular if and only if there exists a point  $w_0 \in W$  such that*

$$\lim_{n \rightarrow \infty} \text{dist}(\sigma_n(w), w_0) = 0 \quad \text{for all } w \in W \quad (2.34)$$

where  $\text{dist}(A, w) := \inf_{w' \in A} |w' - w|$  for  $A \subset W$ .

- (ii) *For  $(\sigma_n)$  in (2.33), the following holds:*

$$\sigma_{m+n}(w) = \overline{\bigcup_{w' \in \sigma_m(w)} \sigma_n(w')}, \quad (2.35)$$

for all  $m, n \in \mathbb{N}_+$ ,  $w \in W$ , where the overline means the topological closure in  $W$ .

**Proof.** (i) See Theorem 3.3.31, p.116, of [9].

(ii) By assumption,  $W$  is a compact metric subspace of  $\mathbb{R}$ . Hence the  $Q$  in (2.10) is “continuous” in the sense of Definition 3.3.1 in [9]. Therefore we can apply Lemma 3.3.32, p. 117 of [9] and the statement holds.  $\square$

**Definition 2.13.** *Let  $\{(W, \mathcal{W}), X, u, P\}$  be an RSCC and let  $P_r$  be as in (2.7).*

- (i) *For  $w \in W$ ,  $n, r \in \mathbb{N}_+$ , and  $A \subset X^r$ , define*

$$P_r^n(w, A) := P_{r+n-1}(w, X^{n-1} \times A), \quad (2.36)$$

with the convention  $X^0 \times A := A$ .

- (ii) ([9], Definition 2.1.4) *An RSCC  $\{(W, \mathcal{W}), X, u, P\}$  is said to be uniformly ergodic if for any  $r \in \mathbb{N}_+$ , there exists a probability measure  $P_r^\infty$  on  $(X^r, \mathcal{P}(X^r))$  such that the sequence  $\{P_r^n(w, A) : n \geq 1\}$  in (2.36) converges uniformly to  $P_r^\infty(A)$ , that is, the following sequence  $(\varepsilon_n)_{n \in \mathbb{N}_+}$  goes to 0 when  $n \rightarrow \infty$ :*

$$\varepsilon_n := \sup\{|P_r^n(w, A) - P_r^\infty(A)| : w \in W, r \in \mathbb{N}_+, A \subset X^r\}. \quad (2.37)$$

**Theorem 2.14.** *Under the Assumption (B), let  $\{(W, \mathcal{W}), X, u, P\}$  be an RSCC with contraction such that  $W$  is compact. Assume that  $\{(W, \mathcal{W}), X, u, P\}$  has a regular associated Markov chain. Then the following holds:*

(i)  $\{(W, \mathcal{W}), X, u, P\}$  is uniformly ergodic.

(ii) Let  $Q^\infty$  be as in Remark 2.10(iii). Then the probability measure  $P_r^\infty$  in (2.37) is given as follows:

$$P_r^\infty(A) = \int_W P_r(w, A) dQ^\infty(w), \quad A \in \mathcal{P}(X^r) \quad (2.38)$$

where  $P_r$  is as in (2.7).

**Proof.** (i) From Theorem 3.4.5, p.125 in [9], the statement holds. (ii) By (i), conditions in Remark 2.10(ii) and (iii) are satisfied by assumption. Hence we obtain  $Q^\infty$  on  $(W, \mathcal{W})$  in (2.29). From Theorem 3.4.5, p.125 in [9], the statement holds.  $\square$

### 3 Proof of Theorem 1.1

In this section, we prove Theorem 1.1. General results in Section 2 will be applied to  $N$ -continued fraction expansions.

#### 3.1 RSCC associated with $N$ -continued fraction expansion

Fix an integer  $N \geq 1$ . We introduce a random system with complete connections (=RSCC)  $\{(I, \mathcal{B}_I), \Lambda, u, P\}$  as follows:

$$\left\{ \begin{array}{ll} u : I \times \Lambda \rightarrow I; & u(x, i) := u_{N,i}(x) = \frac{N}{x+i}, \\ P : I \times \Lambda \rightarrow I; & P(x, i) := V_{N,i}(x) = \frac{x+N}{(x+i)(x+i+1)}, \\ \Lambda = \{N, N+1, \dots\}. \end{array} \right. \quad (3.1)$$

By the definition of  $P$  and using the partial fraction decomposition,  $\sum_{i \in \Lambda} P(x, i) = 1$ . By (2.10),  $Q(x, B) = \sum_{i \in X(x, B)} V_{N,i}(x)$  for  $(x, B) \in I \times \mathcal{B}_I$  where  $X(x, B) := \{i \in \Lambda : u_{N,i}(x) \in B\}$ . Let  $U$  and  $Q^{(n)}$  be as in

(2.21) and (2.14), respectively. By definition,  $\{(I, \mathcal{B}_I), \Lambda, u, P\}$  satisfies the Assumption (B) in Section 2.4.

For the dynamical system  $(I, \mathcal{B}_I, T_N)$  in Section 1 and a given probability measure  $\mu$  on  $(I, \mathcal{B}_I)$ , the ergodic behavior of the RSCC in (3.1) allows us to find the limiting Gauss-Kuzmin distribution  $F$  with respect to  $(T_N, \mu)$ :

$$F(x) := \lim_{n \rightarrow \infty} \mu(T_N^n < x), \quad x \in I \quad (3.2)$$

and the invariant measure induced by  $F$ .

**Lemma 3.1.**  $\{(I, \mathcal{B}_I), \Lambda, u, P\}$  is an RSCC with contraction.

**Proof.** We have

$$\begin{aligned} \frac{d}{dx} u(x, i) &= \frac{d}{dx} u_{N,i}(x) = -\frac{N}{(x+i)^2}, \\ \frac{d}{dx} P(x, i) &= \frac{i^2 + i - 2Ni - (x^2 + 2Nx + N)}{(x+i)^2(x+i+1)^2} \end{aligned} \quad (3.3)$$

for any  $x \in I$  and  $i \in \Lambda$ . Thus,

$$\sup_{x \in I} \left| \frac{d}{dx} u(x, i) \right| \leq \frac{N}{i^2}, \quad i \in \Lambda \quad (3.4)$$

$$\sup_{x \in I} \left| \frac{d}{dx} P(x, i) \right| < \infty. \quad (3.5)$$

Hence the requirements of Definition 2.11 are fulfilled.  $\square$

**Lemma 3.2.**  $\{(I, \mathcal{B}_I), \Lambda, u, P\}$  has a regular associated Markov chain.

**Proof.** By Theorem 2.12(i), it is equivalent to that

$$(\Gamma) : \left\{ \begin{array}{l} \text{there exists a point } x^* \in I \text{ such that} \\ \lim_{n \rightarrow \infty} \text{dist}(\sigma_n(x), x^*) = 0 \text{ for all } x \in I \end{array} \right\}$$

where we remark that  $W = I = [0, 1]$  in this case. Hence we show the condition  $(\Gamma)$  as follows.

Fix  $x \in I$ . Let us define the sequence  $(x_n)_{n \geq 0}$  in  $I$ , recursively by  $x_0 := x$ ,  $x_{n+1} := \frac{N}{x_n + N}$  for  $n \geq 1$ . Clearly  $x_{n+1} \in \sigma_1(x_n)$  and therefore Theorem 2.12(ii) and an induction argument lead us to the conclusion that  $x_n \in \sigma_n(x)$  for  $n \in \mathbb{N}_+$ . But,  $\lim_{n \rightarrow \infty} x_n = x^* = \frac{-N + \sqrt{N^2 + 4N}}{2}$  for any  $x \in I$ . Hence  $\text{dist}(\sigma_n(x), x^*) \leq |x_n - x^*| \rightarrow 0$  as  $n \rightarrow \infty$ . Hence we find  $x^* := \frac{-N + \sqrt{N^2 + 4N}}{2}$  in the condition  $(\Gamma)$ .  $\square$

**Proposition 3.3.**  $\{(I, \mathcal{B}_I), \Lambda, u, P\}$  is uniformly ergodic.

**Proof.** In order to apply Theorem 2.14 to the RSCC  $\{(I, \mathcal{B}_I), \Lambda, u, P\}$  in (3.1), we check assumptions in Theorem 2.14. By definition,  $I$  is compact. By Lemma 3.1,  $\{(I, \mathcal{B}_I), \Lambda, u, P\}$  is an RSCC with contraction. By Lemma 3.2,  $\{(I, \mathcal{B}_I), \Lambda, u, P\}$  is regular. Hence all assumptions in Theorem 2.14 are satisfied. Hence the statement holds.  $\square$

Let  $L(I)$  denote the Banach space of all complex-valued Lipschitz continuous functions on  $I$ . The regularity of  $U$  in (1.4) with respect to  $L(I)$  follows from Theorem 2.12. Moreover, the sequence  $\{Q^{(n)}(\cdot, \cdot) : n \geq 1\}$  in (2.14) converges uniformly to a probability measure  $Q^{(\infty)}$  on  $(I, \mathcal{B}_I)$  and that there exist two positive constants  $q < 1$  and  $k$  such that

$$\|U_n f - U_\infty f\|_L \leq k q^n \|f\|_L, \quad n \in \mathbb{N}_+, f \in L(I) \quad (3.6)$$

where

$$U_n : L(I) \rightarrow L(I); \quad \{U_n f\}(x) = \int_I f(y) dQ_x^{(n)}(y), \quad (3.7)$$

$$U_\infty : L(I) \rightarrow \mathbb{C}; \quad U_\infty f = \int_I f(y) dQ^{(\infty)}(y) \quad (3.8)$$

with  $Q_x^{(n)}(B) := Q^{(n)}(x, B)$  in (2.14).

**Proposition 3.4.** The probability  $Q^{(\infty)}$  is the invariant probability measure of the transformation  $T_N$ .

**Proof.** For  $G_N$  in Section 1 and  $Q$  in (2.10), and on account of the uniqueness of  $Q^{(\infty)}$  we have to show that

$$\int_0^1 Q(x, B) dG_N(x) = G_N(B), \quad B \in \mathcal{B}_I. \quad (3.9)$$

Since the intervals  $[0, u) \subset [0, 1)$  generate  $\mathcal{B}_I$ , it is sufficient to show the equation (3.9) just for  $B = [0, u)$ ,  $0 < u \leq 1$ . Let  $E(x, N) = \lfloor \frac{N}{u} - x \rfloor + 1$ . Since  $Q(x, B) = \sum_{i \in X(x, B)} V_{N, i}(x)$  for  $(x, B) \in I \times \mathcal{B}_I$  where  $X(x, B) := \{i \in \mathbb{N} : u_{N, i}(x) \in B\}$ , we have

$$\begin{aligned} Q_m(x, [0, u)) &= \sum_{\{i \in \mathbb{N} : 0 \leq u_i(x) < u\}} V_{N, i}(x) = \sum_{i \geq E(x, N)} V_{N, i}(x) \\ &= \frac{N - E(x, N)}{x + E(x, N)}. \end{aligned} \quad (3.10)$$



Thus,

$$\int_0^1 Q(x, [0, u)) dG_N(x) = \frac{1}{\log\{(N+1)/N\}} \log \frac{x+N}{N} = G_N([0, u)). \quad (3.11)$$

Hence the statement holds.  $\square$

### 3.2 Proof of Theorem 1.1

By using Proposition 3.3, we prove Theorem 1.1.

**Proof of Theorem 1.1** Fix an integer  $N \geq 1$  and let  $G_N$  be as in Section 1. By (1.5), we have

$$\mu(T_N^{-n}(A)) = \int_A \{U^n f_0\}(x) dG_N(x) \quad \text{for any } n \in \mathbb{N}_+, A \in \mathcal{B}_I \quad (3.12)$$

where  $f_0(x) = (x+N)(d\mu/d\lambda)(x)$  for  $x \in I$ . If  $d\mu/d\lambda \in L(I)$ , then  $f_0 \in L(I)$  and by (3.8) we have

$$\begin{aligned} U^{(\infty)} f_0 &= \int_I f_0(x) Q^{(\infty)}(dx) = \int_I f_0(x) G_N(dx) \\ &= \int_I (d\mu/d\lambda)(x) dx = \frac{1}{\log\{(N+1)/N\}} \log \frac{x+N}{N}. \end{aligned} \quad (3.13)$$

Taking into account (3.6), there exist two constants  $q < 1$  and  $k$  such that

$$\|U^n f_0 - U^{(\infty)} f_0\|_L \leq kq^n \|f_0\|_L, \quad n \in \mathbb{N}_+. \quad (3.14)$$

Furthermore, consider the Banach space  $C(I)$  of all real-valued continuous functions on  $I$  with the norm  $\|f\| := \sup_{x \in I} |f(x)|$ . Since  $L(I)$  is a dense subspace of  $C(I)$  we have

$$\lim_{n \rightarrow \infty} \|(U^n - U^{(\infty)})f\| = 0 \quad \text{for all } f \in C(I). \quad (3.15)$$

Therefore, (3.15) is valid for a measurable function  $f_0$  which is  $Q^{(\infty)}$ -almost surely continuous, that is, for a Riemann-integrable function  $f$ . Thus, we have

$$\lim_{n \rightarrow \infty} \mu(G_N^n < x) = \lim_{n \rightarrow \infty} \int_0^x \{U^n f_0\}(u) G_N(du) \quad (3.16)$$

$$= \frac{1}{\log\{(N+1)/N\}} \log \frac{x+N}{N} \int_0^x G_N(du) \quad (3.17)$$

$$= \frac{1}{\log\{(N+1)/N\}} \log \frac{x+N}{N}. \quad (3.18)$$

Hence (1.6) is proved.  $\square$

## References

- [1] Barnsley, M., Elton, J., *A New Class of Markov Processes for Image Encoding*, Adv. in Appl. Probab. **20** (1988), 14-32.
- [2] Barnsley, M., Demko, S., Elton, J., Gerinomo, J., *Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities*, Ann. Inst. H. Poincaré, Probab. Statist **24**(3) (1988), 367-394.
- [3] Burger, E. B., Gell-Redman, J., Kravitz, R., Walton, D., Yates, N., *Shrinking the period lengths of continued fractions while still capturing convergents*, J. Number Theory, **128**(1) (2008) 144-153.
- [4] Dajani, K., Kraaikamp, C., van der Wekken, N., *Ergodicity of  $N$ -continued fraction expansions*, J. Number Theory, **133**(9) (2013) 3183-3204.
- [5] Doeblin, W., Fortet, R., *Sur des chaînes à liaisons complètes*, Bull. Soc. Math. France **65** (1937), 132-148.
- [6] Durrett, R., *Probability theory: Theory and examples 3rd ed.*, Thomson Brooks/Cole 2005.
- [7] Harris, T.E., *On chains of infinite order*, Pacific J. Math. **5**(1955), 707-724.
- [8] Iosifescu, M., *Random systems with complete connections with an arbitrary set of states*, Rev. Roumaine Math. Pures Appl. **8** (1963), 611-645.
- [9] Iosifescu, M., Grigorescu, S., *Dependence With Complete Connections and its Applications*, Cambridge Tracts in Mathematics **96**, Cambridge Univ.Press, Cambridge, 2nd edition, 2009.
- [10] Iosifescu, M., Theodorescu, R., *Random Processes and Learning*, Springer-Verlag, Berlin, 1969.
- [11] Kalpazidou, S., *On a problem of Gauss-Kuzmin type for continued fraction with odd partial quotients*, Pacific J. Math. **123** (1) (1986), 103-114.

- [12] Karlin, S., *Some random walks arising in learning models. I*, Pacific J. Math. **3** (4) (1953), 725-756.
- [13] Kawamura, K., *Automata computation of branching laws for endomorphisms of Cuntz algebras*, Int. J. Alg. Comput. **17**(7) (2007), 1389-1409.
- [14] Kawamura, K., *A generalization of de Bruijn graphs and classification of endomorphisms of Cuntz algebras by graph invariants*, Semigroup Forum **81** (2010), 405-423.
- [15] Kawamura, K., Hayashi, Y., Lascu, D., *Continued fraction expansions and permutative representations of the Cuntz algebra  $\mathcal{O}_\infty$* , J. Number Theory **129** (12) (2009), 3069-3080.
- [16] Lascu, D., *On a Gauss-Kuzmin-type problem for a family of continued fraction expansions*, J. Number Theory **133**(7) (2013), 2153-2181.
- [17] Lascu, D., *Metric properties of  $N$ -continued fractions*, manuscript.
- [18] Lothaire, M., *Combinatorics on words*, Cambridge University Press, 1997.
- [19] Mauldin, R.D., Urbański, M., *The doubling property of conformal measures of infinite iterated function systems*, J. Number Theory **102** (1) (2003), 23-40.
- [20] Mealy, G.J., *A method for synthesizing sequential circuits*, Bell Syst. Tech. J. **34**(5)(1955) 1045-1079.
- [21] Norman, E., *Markov Processes and Learning Models*, Academic Press, New York, 1972.
- [22] Onicescu, O., Mihoc, Gh., *Sur les chaînes de variables statistiques*, Bull. Sci. Math. **59** (1935), 174-192.
- [23] Sebe, G.I., *On convergence rate in the Gauss-Kuzmin problem for grotesque continued fractions*, Monatsh. Math. **133** (2001), 241-254.
- [24] Sebe, G.I., *A Gauss-Kuzmin theorem for the Rosen fractions*, J. Théor. Nombres Bordeaux **14** (2) (2002), 667-682.
- [25] Sebe, G.I., Lascu, D., *A Gauss-Kuzmin theorem and related questions for  $\theta$ -expansions*, Journal of Function Spaces, vol. **2014** (2014), 12 pages.